AD________________

GRANT NUMBER DAMD17-96-1-6153

TITLE: A Well Tempered Mammographic Display

PRINCIPAL INVESTIGATOR: Harold Kundel, M.D.

CONTRACTING ORGANIZATION: University of Pennsylvania
Philadelphia, PA 19104

REPORT DATE: September 1998

TYPE OF REPORT: Final

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
September 1998

3. REPORT TYPE AND DATES COVERED
Final (1 Sep 96 - 31 Aug 98)

4. TITLE AND SUBTITLE
A Well Tempered Mammographic Display

5. FUNDING NUMBERS
DAMD17-96-1-6153

6. AUTHOR(S)
Harold Kundel, M.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Pennsylvania
Philadelphia, PA 19104

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

19981231 045

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200*

The objective was the development and testing of a simple softcopy display for digital mammography that was matched to human contrast sensitivity at ambient light levels.

A calibration procedure was developed that derived a display function for the mammogram display based upon equalizing the just noticeable difference in contrast over the intensity range of the monitor at ambient illumination. The maximum and minimum pixel intensity determined for each breast image was used to set the range of the pixel driving value. The two procedures produce a perceptually tempered display. A single non-linear adjustment controls the displayed contrast. A rove-zoom feature that simultaneously coupled the location of features in the cranio-caudal and medio-lateral oblique views was developed.

The time to detection and the detection accuracy using ROC analysis was compared for two radiologists reading mammograms using three display functions: 1) a function specifically set for each image that acted as an ideal performance standard, 2) standard linear contrast controls and 3) the perceptually tempered display. There was no difference in accuracy between the standard linear contrast controls and the perceptually tempered display. Both display functions were as accurate as the performance standard but they both incurred slightly more decision time.

14. SUBJECT TERMS Breast Cancer

15. NUMBER OF PAGES
14

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

HK Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Hanna Krendel 9/28/98

PI - Signature Date

TABLE OF CONTENTS

5. Introduction

5.1. Purpose of the Project

The long term goal is to improve the detection of masses on digital mammograms by using a softcopy display that will optimize the coupling of human vision to the display without excessive human-machine interaction. This coupling is called a "perceptually tempered" display.

5.2. Scope of this Project

The program involved the development of computer software in the IDL language (Research Systems, Inc., Boulder, CO) for producing a perceptually tempered display in which the displayed contrast on a mammogram is matched to the contrast sensitivity of the reader at the ambient illumination. The software is evaluated subjectively by having a mammographer use the display station. This is followed by testing detection using hybrid images consisting of simulated masses embedded in sections of real mammograms. The final test consists of a difficult set of real mammograms read by radiologists. The test includes speed and accuracy as measured by receiver operating characteristic (ROC) analysis. The software was first developed to operate on the entire image and then selectively within a roving window on the display.

5.3. Background

For the foreseeable future, mammograms will be the primary screening tools for breast cancer and they will be read by human beings. Whenever images are read by human beings there is an associated error rate. Although the true diagnostic accuracy for screening examinations is unknown, false negatives (misses) have been estimated to range between 11 and 25 % and false positives have been estimated to range between 10 and 35 % . These figures indicate that there is room for improvement in the readers.

Digital detectors will be a reality in the near future making softcopy readings feasible. At present, softcopy is less effective than hardcopy because, as compared with currently available films, monitors do not have sufficient intensity range, do not have sufficient resolution and are too noisy. Although bright (1000 cd/m^2), high resolution (4000 x 4000 pixel) monitors are feasible, they are too expensive for widespread use. The window-level and zoom-rove functions that are commonly used with cross-section CT and MRI are not sufficient for mammograms. Given the present state of the art, softcopy displays cannot be made that simulate film. Therefore, it is important to develop display mode alternatives to the film-on-lightbox design or cross-section imaging design that can be used with moderate brightness (300 cd/m^2), moderate resolution (2000 x 2000 pixel) monitors.

Johnson et al. [8.1] and Blume et al. [8.2] have suggested that video monitors should be "perceptually standardized" so that equal changes in the pixel gray scale value produce equal changes in the just noticeable difference (JND) of luminance in the image. Although perceptual standardization is a worthwhile way to make monitors adhere to the same standard input-output transfer characteristic, it does not adjust for local contrast variations in the image. In order to match the display of a particular image to the visual system, the image itself must be modified. The usual method is histogram equalization that improves the distribution of gray levels over the display [8.3]. Histogram equalization may not lead to optimal coupling of the display contrast to the contrast sensitivity of the eye because the algorithm values all contrast levels equally while the non-linear visual system is more sensitive to contrast near the adaptation level. Liu and Nodine [8.4] using a model first proposed by Mokrane [8.5] have developed an algorithm that equalize perceived luminance over the image assuming some starting level of adapting luminance. Contrast is modified in the image on the basis of the theoretical threshold-contrast curves of Heinemann [8.6].

This research extends the work of Liu and Nodine [8.4] to include calibration of the individual observer and to incorporate perceptually derived scanning strategies into the search for masses on mammograms.

6. Progress Report

6.1 Technical Objective 1. Show that equalization of perceived contrast within a region of a mammogram will improve the detection of masses.

6.1.1 Aim 1. Produce a large set of simulated mammograms with just detectable masses (detectability index ,d'= 1.0) and backgrounds of additive noise.

6.1.1.1 Background Parenchyma

Mammogram parenchyma patches were chosen from a set of ten digitized mammograms that were selected by a mammographer who judged them to be free of any distinctive abnormal radiological findings suggestive of benign or malignant disease. They were digitized using a Lumisys digitizer to a pixel size of 100 microns with a 12 bit intensity scale. The digitized images were normalized using histogram equalization so that the mean intensity of each displayed image was about 24 cd/m^2. The patches were selected for display by first randomly selecting a mammogram and then randomly selecting a point within the mammogram to locate the center of the patch. The patches were displayed on a 2000 x 2000 pixel monitor (Tektronix , Beaverton, OR) with a dynamic range of 0 - 174 cd/m^2. Prior to each reading session the monitor was calibrated using a J16 photometer (Tektronix, Beaverton, OR).

6.1.1.2 Simulated Masses

Masses were simulated as projections of a 1 cm spherical object of soft tissue x-ray attenuation embedded in a homogeneous medium. A blur factor was added for focal spot and geometrical unsharpness. The simulator computed the intensity distribution of the masses using 5 levels of intensity that when added to the backgrounds produced contrasts of about .01, .02, .03, .04, and .05 for the gaussian noise and .02, .04, .06, .08, and .1 for the mammogram. The gaussian noise was computer generated and had a mean intensity of 85 cd/m^2 and a standard deviation of 14 cd/m^2. The displayed luminance was used to calculate displayed contrast as

$$\text{Displayed Contrast} = \frac{\text{Luminance with Mass on Background} - \text{Luminance of Background Only}}{\text{Luminance with Mass on Background} + \text{Luminance of Background Only}}$$

6.1.2 Preliminary Detection Experiment: Mammogram Background v Gaussian Noise

Three volunteer subjects who were not radiologists participated in 20 viewing sessions each consisting of 200 trials. A two alternative forced choice (2AFC) method was used. The subjects received feedback about the correct location of the mass after each trial. Prior to the experimental sessions each subject had about 200 practice trials. Sessions using a gaussian background were mixed randomly with the sessions using a mammogram background. The 5 contrast levels were randomized over trials. At the completion of the study, each subject saw each contrast level 200 times for a total of 4000 trials per subject. The index of detectability, d', was calculated for each subject at each contrast level [8.7]. A linear regression analysis using d' and average mass contrast was performed for each subject in each condition. The results are shown in Table 1 and Figure 1.

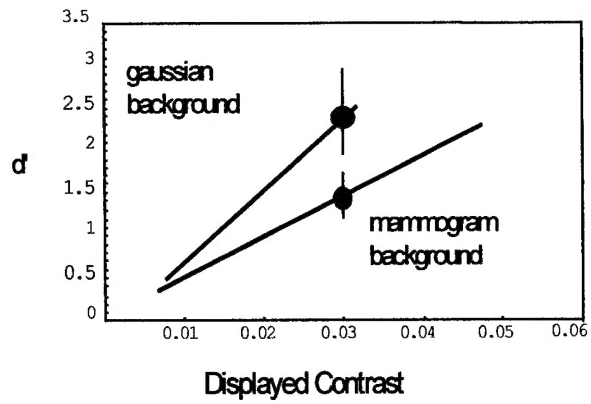

Figure 1. A Comparison of the regression lines obtained in an experiment in which 3 readers had to report 1 cm. masses embedded in a background of either gaussian noise or normal mammogram parenchyma. The displayed contrast of the mass is plotted against the index of detectability, d'. The single data point was included to show the 95% confidence range. These data were collected in a preliminary experiment to show that the mammogram background made mass detection more difficult and that simulated masses with a d' ranging from 1 to 2 could be synthesized reliably.

6.1.3 Aim 2: Apply the contrast equalization algorithm of Liu and Nodine to circular regions of the image test set with mammogram backgrounds. Compare performance with and without the algorithm at a variety of adaptation levels.

6.1.3.1 Approximation of the Contrast Sensitivity Curve by a Parabola.

The algorithm of Liu and Nodine required advanced information about adaptation level and was computationally intensive. We simplified the algorithm by assuming that the relationship could be approximated by a parabola. (see Figure 4). Integration of the parabola determined the shape of the look-up table. We needed to find the shape of the parabola for each observer and the range of the parabola for each image. For example, a dark image from a fatty breast would have a different range than a bright image from a dense breast. The look-up table would be different for each image.

A two step procedure was developed. First, the shape of the parabola was determined using a Minimal Detectable Contrast (MDC) correction. Second, the breast was located in the image and the pixels were sampled to determine the range (Minimum and Maximum parenchymal pixel value). These procedures are described below.

6.1.3.2 Minimal Detectable Contrast (MDC) Look-up Table Correction Process.

An MDC test pattern is displayed to each observer prior to viewing session. The MDC test pattern consists of 8 horizontal bands of increasing intensity (figure 2). Each band contains 8 circular targets of increasing contrast. The observer's task is to choose the "least detectable" target in each band. The contrast of each indicated target is used to fit a 2^{nd} degree equation, where the independent variable is the driving level of the intensity band and the dependent variable is the contrast of the observer indicated target (figure 3) in pixel driving level units. This curve is fitted by a parabola, where the least visible targets in the most dark and most light bands are necessarily of higher pixel driving level contrast. The resultant best fit curve can be seen as a measure of the minimum detectable contrast requirements under the existing viewing conditions.

The resultant best-fit curve is integrated and normalized to the available display intensity range to yield a continuous, non-linear lookup table, resulting in a lookup table that boosts contrast in the intensity bands that require higher contrast for detection of low contrast targets. On the other hand, because there are a limited number of available intensities, the targets that reside in the remaining regions are of lower contrast after the MDC correction is applied. The MDC lookup table is designed to equalize the detectability of equal contrast (pixel driving level) targets, regardless of the regional mean pixel intensity surrounding the targets.

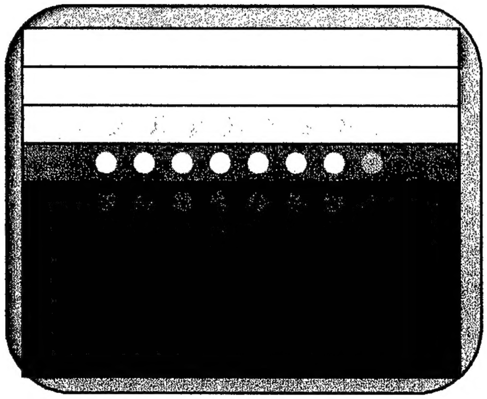

Figure 2: MDC Test Pattern

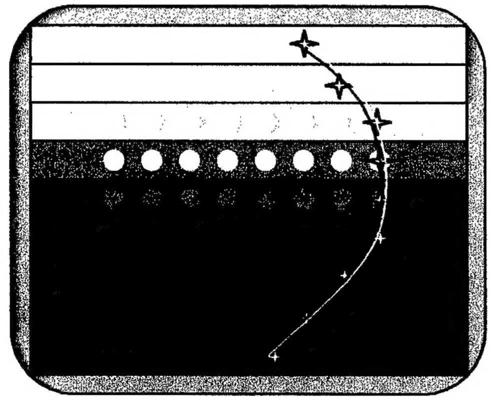

Figure 3: Observer indicated these points as minimally detectable. The best fit 2^{nd} degree equation is fitted.

6.1.2.3 Matching the Look-up Table to the Pixel Intensity of the Mammogram.

As each case is displayed, the MDC corrected lookup table is modified so that the endpoints of the table anchor at the minimum and maximum sampled intensities of the image (Figure 6). The images are sampled over a region that includes breast tissue out to the skinline, but that excludes the extremes of pixel driving levels due to lead markers, labels and film/cassette edge artifacts. The selected area for intensity analysis is customized for each image by an edge detection procedure operating on binary version of the median filtered image. Image intensities are then sampled along evenly spaced lines as shown in figure 5. Matching the lookup table to pixel values in the breast and out to the skinline is done to more efficiently visualize the appropriate tissue densities for diagnostic information.

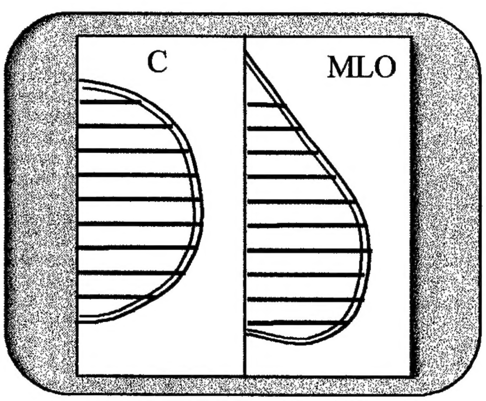

Figure 5: The intensities are sampled in regions are selected to avoid non-tissue related regions such as labels and markers, but does sample to beyond the skinline

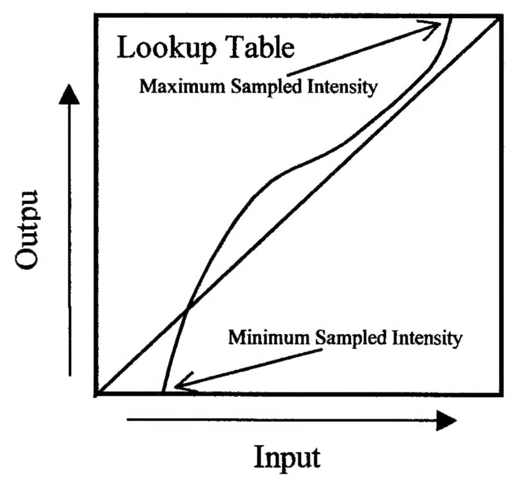

Figure 6: The resulting MDC lookup table is optimized to appropriate pixel intensities in the breast images.

Observers are able to use a single slider to adjust the severity (gamma) of the MDC lookup table. The slider can smoothly adjust the gamma from an almost linear lookup table up to a maximum MDC setting.

6.1.3 Aim 3: Produce a set of simulated mammograms with just detectable masses (d' =1) and realistic backgrounds. Compare performance with and without the algorithm at a variety of adaptation levels.

The experimental effort to deal with adaptation levels singly (3 observers x 200 trials x 8 levels) was beyond our resources (too many trials, too much time, too hard to get unpaid volunteers to read for hours at a time) and it was decided to use instead a test set of 75 real mammograms (25 cancers, 25 benign lesions and 25 difficult normals) digitized to 100 microns using a Lumisys digitizer. A number of readers had already viewed the mammogram test set and we knew that the average d' was about 1.0, which corresponds to an area (Az) under the receiver operating characteristic (ROC) curve of .76 .

6.1.3.1 Observer Experiment

Two radiologists viewed the mammogram test set which consisted of 75 cases where each cases consisted of a cranio-caudal (CC) and medio-lateral oblique (MLO) view. The contralateral mammogram and prior views were not available. They were asked to localize each area on the mammogram that was suspicious for malignancy and give their confidence that it was malignant (1 to 5). Three viewing conditions were used. No feedback was provided. At least 6 weeks elapsed between readings. The three conditions were:

1) images with a look-up table preset by an observer with full knowledge of the pathology on each image;
2) images with a standard window and level adjustment as well as a zoom-and rove window directed by the mouse;
3) images with the perceptually tempered look-up table and a full resolution perceptually tempered roving window directed by the mouse.

Condition 1 represented the best image that was available. This is called the "visual preset" or the "performance standard". Condition 2 represented the type of softcopy display currently available with digital mammography systems. Condition 3 represented the perceptually tempered display with a window and without background suppression.

6.1.3.2 Ambient Room Conditions:

An observer's sensitivity to low contrast targets will be affected by luminance settings of the monitor, ambient room illumination level and back light reflections off the display screen surface. In an attempt to control these variables, we have situated the display surface to minimize back light reflections. The room illumination level is set to .15 foot-candles using a rheostat controlled quartz-halogen torch lamp that is place 10 feet to the side of the display monitor. The room illumination is measured from a point on top of the monitor casing using a Tektronix J17 photometer with a J6511 illuminance probe. The monitor luminance settings are set by the manufacturer to yield a dark luminance of 1.7 cd/m^2 and a maximum brightness of 346 cd/m^2. The 21" Orwin (model D2100L, Clinton Electronics Corporation, 6701 Clinton Road, Rockford, IL 61111) gray scale monitor self-calibrates ½ hour after being turned on and every 24 hours of operation thereafter.

6.1.3.3 Measurements

The time required to make a diagnosis and the diagnostic accuracy as defined by the area under the ROC curve were measured.

The results shown in Table 2 and 3 indicate that the perceptually tempered display made no difference in performance.

6.2 Technological Objective #2 and #3

TE#2. Show that equalizing perceived contrast in a local region together with suppression of background intense variation is more effective than equalizing perceived contrast over the entire image.

TE#3. Determine if masses of unknown location can be detected more effectively using the contrast equalization algorithm in a roving window.

Objective #2 and Objective #3 both require the development of a window that can be moved over the image. In our discussions with the mammographers, there was some controversy about whether the CC and the MLO views should be viewed together or separately. We decided to view them together with a coordinated window, that is, when the window was on the CC view, the corresponding area on the MLO view should be highlighted and vice versa. This involved a considerable programming effort and a verification study using calcifications visible on both CC and MLO views to show that the two windows were really coordinated.

6.2.1 A Coordinated Locator Window for Two Projections (CC and Oblique) of the Same Breast.

When a mammogram series is made on a patient, each breast is compressed between two plates to minimize the tissue thickness in a direction parallel to the x-ray beam direction. This has the effect of spreading the breast tissue in the two remaining dimensions. Subsequent views on a breast require the breast to be de-compressed, rotated (relative to the compression plates) and re-compressed in a different direction. It can be difficult to locate the equivalent piece of breast tissue in alternate views of the same breast due to the combined effects of perspective changes along with the compression of tissue along different axes. As a potential decision aid, we have developed a method of calculating a region of interest on a second view (eg: MLO) of a breast when a point or region of interest on a first view (eg: CC) is selected. This is accomplished by constructing a mathematical volumetric model of breast, and mathematically decompressing, rotating and recompressing the model to determine where a volume of breast tissue in one view might fall in the alternate view. We model the compression and decompression process using a linear approximator (scaling operator), namely, a foreshortening of the volume model along the beam direction axis and an expansion in the remaining two dimensions. The model is then rotated around the axes that approximate the actual breast rotation, then recompressed. An example of this sequence using a sphere for demonstrative purposes is shown in figure 7.

In order to accomplish the transformation of a volume from one view to another, we need some basic spatial information about the breast projections and geometry. The parameters that we require are the location of the two axes of rotation and some measure of the compression used when the mammogram is taken. The axes are selected by an automatic procedure that works as follows:

- The breast tissue edges are found using an edge detection process on a median filtered image.
- The plate rotation axis is approximated by finding the line of maximum thickness of the breast outline in the x-dimension on the CC view, as shown in figure x. This should occur close to both the nipple shadow and the center of the y-dimension of the film if the breast was properly centered in the compression plates. The compression plate rotation angle is written into the patient label on the corner of the film.
- The patient is often rotated between 10-20 degrees by being told to rotate her shoulders. This is done to expose underarm musculature. When this angle is not recorded it must be approximated. The axis itself is approximated as the same distance from the nipple as the plate rotation axis.

The compression parameters are also estimated using a ratio that approximates the change in the thickness of the breast tissue when compressed. The critical value is related to the relative change in the

length of a line element within the breast tissue along the two compression axes. For the purposes of our initial locator procedure, we set this to a constant, assuming that similar levels of compression are used over exams. We suspect that the compression varies from one patient to the next in clinical exams.

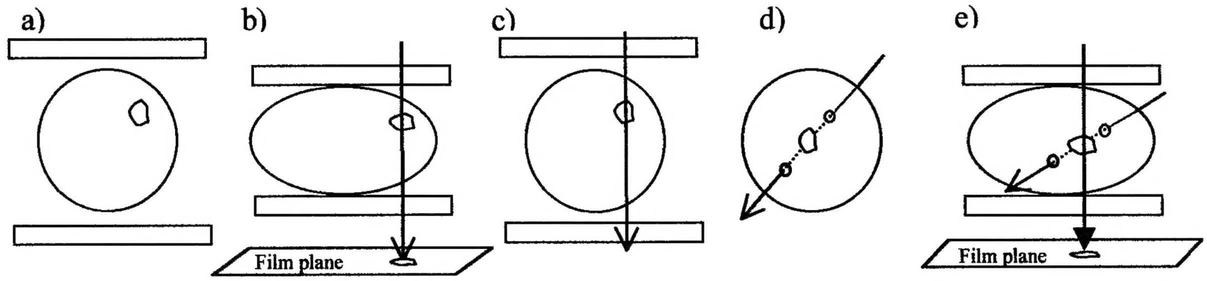

Figure 7: a) Original volumetric object with potential target site. b) Compressed volume showing projection beam through target to film plane (CC view) c) Uncompressed volume d) rotate volume relative to compression plates and film plane e) Compressed volume showing projection beam through target to film plane (MLO view).

Since it is not necessary to correlate every point in one view of the breast with every point of the alternate views, we perform the mathematical procedure only on the point or region of interest. A point of interest in the planar CC view represents a line of superposition, seen end-on, in the breast tissue. When this line of superposition is rotated to the MLO view, it is visualized as a line through the tissue. It is important to have the correct plate rotation angle for each view of a breast for the model to yield best accuracy.

There accuracy of the prediction line is dependent upon a number of variables, including knowledge of accurate plate rotation angles, tissue compression ratios and correct patient positioning. The prediction line, therefore, is replaced by a pair of lines that straddle the prediction line, showing the potential error in the positioning of the line. An initial estimate of the standard error was obtained by comparing the prediction lines with known locations for the 25 lesions in our current mammogram data set.

The effect of the coordinated coordinated window on decision accuracy has not been tested. We plan to perform the tests in the future.

6.3 Results

Table 1. The Results of the Linear Regression on the Pooled Data from 3 Subjects. The 95% Confidence Limits of d' at a Contrast of 0.03 were calculated using the Mean Value for the Slope of the Regression Line.

	Slope		d' at .03 Contrast	
Background	Mean	95% CI	Mean	95% CI
Gaussian	88	65 - 110	2.4	1.9 - 2.9
Mammogram	49	41 - 56	1.5	1.2 - 1.7

Table 2. The Area Under the ROC Curve as Mean and (St. Dev) for 2 Radiologists Reading Mammograms From a Softcopy Display Using 3 Different Display Functions.

	Display Function		
	Visual Preset	Standard	Perceptually Tempered
Radiologist 1	.72 (.06)	.63 (.08)	.75 (.06)
Radiologist 2	.75 (.07)	.86 (.04)	.87 (.04)

Table 3. The Time in Seconds Required for 2 Radiologists to Make a Decision About the Presence of a Cancer on a Mammograms Using 3 Different Softcopy Display Functions.

	Display Function		
	Visual Preset	Standard	Perceptually Tempered
Radiologist 1	37	46	50
Radiologist 2	36	50	73

7. Conclusions

7.1 Summary

7.1.1 The speed and accuracy of the perceptually tempered display display function is equal to the standard linear display function when used on a moderately bright monitor (300 cd/sqm).

7.1.2 The issue of the diagnostic accuracy of local as opposed to global image processing has not been resolved by this study.

7.2 Discussion

The original intent of this project was to approach systematically the problem of improving the detection of masses in mammograms by starting with a model system consisting of hybrid mammograms - simulated masses on real backgrounds - and progressing to real mammograms. It was based upon the assumption that correcting the display function for visual adaptation at the dark and light end of the gray scale would improve the performance. However, gains in contrast sensitivity at the extremes of the image (white and black parts) are made at the expense of loss in the center (mid grays). The mammographer really needs a way to enhance contrast locally without losing the contextual information from the rest of the image. We are currently working on a method to automate this function. The observer's preferred contrast settings used to diagnose 450 image trials will be used to study this issue.

8. References

8.1. Johnston RE, Zimmerman JB, Rogers DC, and Pizer SM. Perceptual standardization. SPIE Medical Imaging, 1985; 536:44-49.

8.2. Blume H, Roehrig H, Browne M, Ji TL. Comparison of the physical performance of high resolution CRT displays and films recorded by laser image printers and displayed on light-boxes and the need for a display standard. Proc SPIE 1990;1232:97-114.

8.3. Pizer SM et al.. Adaptive histogram equalization and its variations:, Comput Vision Graphics Image Process., 1987; 39:355-368.

8.4. Liu H and Nodine CF. A generalized image contrast enhancement technique based on the Heinemann contrast discrimination model. J. Elec. Imaging

8.5. Mokrane A. A new image contrast enhancement technique based on a contrast discrimination model. CVGIP: Graphics Models in Image Processing, 1992; 54(2):171-180.

8.6. Heinemann EC. The relation of apparent brightness to the threshold for differences in luminance. J. Exp. Psychol. 1961; 61:389-399.

8.7. Macmillan NA, Creelman CD. Detection Theory: A user's guide. Cambridge, UK: Cambridge University Press, 1991

9. Publications and Abstracts

Kundel HL, Toto LC. Optimizing the visualization of projection images using a photometric, psychometric and image based calibration procedure. In: Kim Y, Mun S, eds. Medical Imaging 1998: Image Display. Proc SPIE, 1998;3335: 528-533.

10. Personnel receiving pay from this effort

Harold L. Kundel, MD	Radiologist	Principal Investigator
Susan G. Orel, MD	Radiologist	Mammographer
Lawrence Toto, BS	Research Specialist	Computer Programmer

11. Appendices NONE